MASTERS THESIS PROPOSAL

Resource-rational Task Decomposition with Theory of Mind

Student:
Sam Acquaviva
samacqua@mit.edu

Advisor:
Prof. Josh Tenenbaum
jbt@mit.edu

Department of Brain and Cognitive Sciences
May 23, 2023

Abstract

Compositionality, the capacity to understand and generate complex ideas by combining simpler ones, is a critical attribute of human intelligence and plays a pivotal role in sophisticated problem-solving tasks. However, prior research on the decomposition of problem-solving tasks has been constrained by several limiting assumptions. Notably, these include neglecting the competence of the agent tasked with problem-solving and failing to effectively interweave subgoal generation with acting towards these subgoals. The proposed thesis aims to redress these deficiencies by pursuing a two-pronged approach. Firstly, I will computationally model and perform a human study to understand how people use Theory of Mind – a causal model of how one's mental state leads to action – to generate subgoals. Secondly, I will extend previous work on resource rational-planning to investigate how people merge planning with action. These human studies will then form the basis for a prompting approach for language models, designed to improve their decomposition ability and augment their performance on complex reasoning tasks.

Contents

1 Introduction

1.1 What is Decomposition?

Compositionality, the ability to determine the meaning of a concept by its sub-parts and their combination, is a fundamental aspect of human intelligence [1, 2, 3, 4, 5, 6]. Compositionality can be viewed as the ability to determine the meaning of $f(g(x))$ when f and g are known functions [7].

Task decomposition, the ability to find sub-problems in a task and use their solutions to solve the original task, requires compositionality and is necessary for efficiently and flexibly addressing the vast array of challenges people must face. This can be viewed as, given a function $h(x)$, finding a set of n functions Z such that $h(x) = z_1(x) \circ z_2(x) ... \circ z_n(x)$, where $z_i \in Z$. When $n = 2$, this reduces to finding f and g such that $h(x) = f(g(x))$.

For example, when planning a route from MIT to Berkeley, one can decompose the problem into first navigating to the airport, then flying to San Francisco, then navigating from San Francisco to Berkeley. In turn, each subtask can be arbitrarily broken down into a set of further subtasks.

1.2 The Necessity of Decomposition

Decomposition is necessitated by the sheer number of potential tasks in the world, which are far too numerous to be approached monolithically. For example, take the task of question-answering. Even if one assumes that any single trivia fact can reasonably be memorized (as potentially evidenced by the high performance of large language models on question-answer benchmarks [8, 9, 10]), the ability to *decompose* the problem into individual questions is still necessary [11]. To answer 2-hop questions (questions such as "Who was the President of the United States the year that LeBron James was born?" which requires using the answer of one question to form the other) quickly becomes infeasible without decomposition, as it requires that each of the n^2 possible combinations of the n facts appear in the training data.

1.3 Reducing the Compositionality Gap

Interestingly, large language models have shown evidence of compositional reasoning ability [12, 13, 14]. However, there remains a significant "compositionality gap" between the individual bits of knowledge that language models possess and their ability to use these pieces of knowledge together to answer new questions (such as multi-hop trivia questions) [15]. While prior work has reduced this gap via prompting, the gap continues to exist, especially on tasks where the decomposition is not clear [16, 17, 18].

One potential method to mitigate this compositionality gap is to study how humans decompose problems. Previous works have indeed studied and modeled how humans decompose problems, but these studies, while providing an interesting normative account of human decisions, have been in limited domains where simple baselines still outperform more complete models. Therefore, this thesis aims to extend current understanding of human problem solving by studying task decomposition in more realistic environments. Concretely, this work will curate new datasets of how humans decompose problems, create new models to explain human performance on these datasets, and use the collected data to inform new strategies for using language models to perform complex reasoning tasks.

2 Related Work

2.1 Planning with Theory of Mind

Theory of Mind is a central construct in cognitive science, denoting our capacity to infer and predict others' thoughts, beliefs, and actions [19]. Prior work argues that, in addition to using Theory of Mind for prediction, people use Theory of Mind when *planning* to change other's actions [20].

This work is limited to planning in simple social settings and does not address task decomposition or subgoal-setting. However, it motivates our thesis work because it shows that in order to predictably change the action of another agent, one must infer their mental states and capabilities. When generating a subgoal (the end-state of one part of a task decomposition), this necessity is clear. In the example of navigating from MIT to Berkeley, one would give different subgoals for an MIT student than they would for a Berkeley student; the MIT student would need less guidance in Boston than a Berkeley student because they are more familiar with the area.

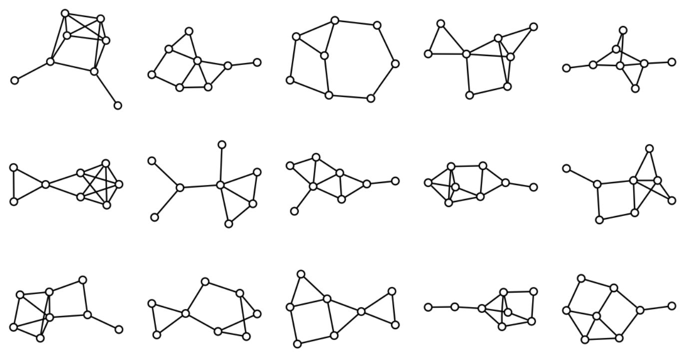

Figure 1: A subset of the graphs presented in [21], where participants chose nodes as subgoals in shortest-path problems.

2.2 Resource Rational Planning and Task Decomposition

Prior work argues that people, when decomposing tasks, are *resource-rational* – meaning that they approximate optimal behavior given their cognitive constraints [21]. While their resource-rational model more closely matches human performance than previous accounts of how humans decompose tasks, it is limited in a few ways. First, it is in a very limited domain – finding the shortest path between two points in small graphs (as shown in Figure 1). Second, it does not outperform a simple baseline called *betweenness centrality*, which (roughly) assumes that people identify bottlenecks in a graph as subgoals. This is likely due to the simplicity of the tasks, where a more complete model is unnecessary to model human performance. Additionally, the provided model is itself a normative account and is not computationally feasible for people.

Other work in the planning literature shows that people are resource-rational when planning in deterministic environments [22]. However, this work does not apply to stochastic environments because their work depends upon the assumption that it is optimal to do all planning before taking action. In reality, when there is uncertainty about the future, one cannot just plan everything and then act out those plans. Instead, people interweave planning and action, balancing the uncertainty of future plans with the sub-optimality of acting before planning out everything.

2.3 Neural Subgoal Proposals

Prior work in the planning domain shows that generating subgoals at different levels of granularity, depending on the complexity of the task at hand, improves performance of hierarchical planning algorithms [23]. However, these subgoals are "blind" to the competence of the agent, and the model requires many expert trajectories or a simulator to be able to generate meaningful subgoals. Because of this limitation, many of the suggested subgoals are unattainable for the agent and must be filtered.

Another technique for incorporating subgoals into problem solving is prompting language models to decompose tasks into smaller parts that are more likely to be solvable. These prompting techniques, such as Chain of Thought (CoT), self-ask, and chain-of-symbol show significant improvements in reasoning performance of language models, but have several limitations [16, 15, 17]. Primarily, they are largely agnostic to insights from cognitive science regarding task decomposition and rely wholly rely on the implicit decomposition ability of the language model. Additionally, these techniques take a very linear approach to problem-solving where the proposed decomposition cannot be changed after it is proposed, which may not always be optimal or reflective of how humans navigate complex problems.

3 Proposed Work

For my thesis work, I will address the limitations outlined in previous works. To do so, I propose to perform two related human studies, model the human behavior from these experiments, and use these data to improve language model task decomposition performance. Specifically, I aim to:

1. Model how people use Theory of Mind to decompose problems, and collect data to validate this model.

2. Model how people use resource-rationality to decompose problems in stochastic environments, and collect data to validate this model.

3. Use these computational accounts of resource-rational task decomposition with Theory of Mind to improve language model performance on tasks requiring decomposition.

3.1 Decomposition with Theory of Mind

The "goodness" of a subgoal depends on the capability of the agent attempting to perform the task. In the example of a student navigating from MIT to Berkeley, the MIT student may need more subgoals in California than they do in Massachusetts, and vice versa. As outlined in Section 2.1 and Section 2.2, prior work in Theory of Mind does not address task decomposition, and prior work in task decomposition does not incorporate Theory of Mind.

In this experiment, we will bridge this gap by collecting data on tasks where the study participant must use Theory of Mind to generate useful subgoals. Specifically, a participant will be shown an agent navigating an environment (like the graphs in Figure 1) with a clear policy, π, trying to reach a goal state, g. Then, after the study participant watches the agent attempt to reach a few different goal states, the agent will be presented with a new goal. The user will be prompted to generate a subgoal, z, for the agent to help the agent navigate towards the goal state, g. Additionally, the participant will be tasked with predicting the agent's next action at different states to probe whether the participant has a good model of the agent's behavior. If humans use Theory of Mind when decomposing problems, the usefulness of their subgoal, z, should increase with their ability to infer the agent's policy, π.

This study offers many benefits over prior work. For one, the necessity of inferring the agent's policy, π, makes it unlikely that simple heuristics will be enough to fully capture human performance. Additionally, it offers a grounded metric of the value of a given task decomposition – the amount of time, t, that it takes the agent to reach the goal state, g. This is in contrast to prior work which only compared the correlation between models of human task decomposition and human performance and found that the betweenness centrality heuristic best fit the human data.

Additionally, it provides a test bed to compare computationally tractable versions of normative accounts of human performance. While previous work was interested in task-agnostic decomposition, which requires reasoning over all possible starting states s and goals g, we are interested in task decompositions that are dependent on the starting state and goal state, which avoids this expensive computation. Additionally, in [21], the reward of a subgoal z is

$$R_{\texttt{Alg}}(s,z) = \sum_{\pi,t} P_{\texttt{Alg}}(\pi,t|s,z)[R(\pi) - t]$$

where $P_{\texttt{Alg}}$ is a distribution over plans, π, and run-times, t, according to the search algorithm `Alg`, and $R(\pi)$ is the reward associated with executing a plan. Intuitively, this equation jointly optimizes task rewards and run-time efficiency. This requires running the search algorithm `Alg` for each considered goal state z and computing the distribution of run-times. Instead, under our model, we would consider the reward of a subgoal z as

$$R(s,z) = \sum_{t} P(t|s,z,\pi')[R(\pi') - t]$$

where π' is an approximation of π dependent on the history of the agent. In this formulation, the distribution over run-times is dependent on the inferred policy π' rather than jointly modeling the run-times and plans. This allows us to separately experiment with computationally feasible approximations to π using Theory of Mind and the distribution over run-times $P(\cdot)$. In summary, our experiment will require models with higher explanatory power, and our proposed model framework will allow us to incorporate Theory of Mind, model computational approximations, and maintain the resource-rationality of the original model.

3.2 Stochastic Task Decomposition

Prior work outlined in Sectin 2.2 shows that people are resource-rational when planning in deterministic environments. However, the real world is stochastic, and it is not always resource rational to do all planning before any action; if the future brings one away from the path they planned, then the effort they spent planning would be wasted. In the MIT to Berkeley navigation example, one might set San Francisco International Airport (SFO) as a subgoal; although there may be other (more optimal) routes that do not go through SFO, buying the ticket allows one to plan the rest of their trip from SFO with less uncertainty. If a flight is delayed or cancelled, then one might still attempt to go to SFO (rather than another route of transportation to the Bay Area) so that they can follow their original plans from SFO.

In this experiment, we will study how people set subgoals as a way to reduce uncertainty in stochastic environments. We will use a stochastic version of the setup used in [22] (show in Figure 2). In this setup, the participant

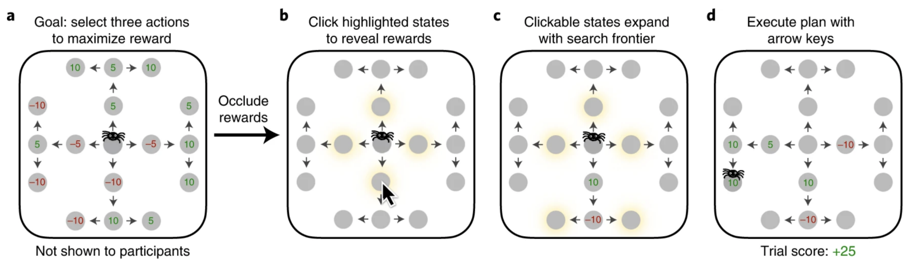

Figure 2: The interface shown to subjects in [22]. In our stochastic version of the experiment, we cannot linearly transition from **b** to **c** to **d**. Instead, the user will be able to dynamically switch between planning (**b** and **c**) and action / subgoal-setting (**d**). Note that, when executing a plan in **d**, the transition in our experiment will be stochastic.

is tasked with maximizing rewards along a path. The participant, in the original study, must first "plan" by revealing the rewards (**b** and **c** in Figure 2), then "act" by using the arrow keys to choose the path (**d** in Figure 2). In our version of the experiment, the actions are stochastic, meaning that if one presses the "Left" arrow key, there is a non-zero probability that they may go another direction. This introduces a new complexity; where previous work could decompose the task into a planning stage and an acting stage, we now need to be able to interweave action and planning, as it is no longer optimal to do all planning before taking any actions.

This study paradigm allows us to investigate a question that was unaddressed in the original work: how do people balance uncertainty with suboptimal behavior? As people are forced to explicitly switch between planning and acting, we can see when people a). stop planning to take action, b). stop acting to plan more, and c). abandon their original plans. This work will be the first to investigate the very non-linear nature of planning in subgoal generation.

3.3 Language Models as Subgoal Generators

Transformer models have been used as subgoal generators to guide search, but this only works with access to expert trajectories or in relatively simple domains, as outlined in Section 2.3. Humans, on the other hand, can generate a more refined search with little to no domain-specific data. Can we augment language models so that, similarly to humans, they can generate useful subgoals in these complex reasoning tasks without access to thousands of training trajectories?

To answer this question, we can run program synthesis with language-model subgoals and compare the number of tasks solved with a fixed compute budget. In these tasks, the subgoals can take many forms: an exact state (e.g. `[True, False, False]`), a state predicate in code (e.g. `[i % 2 == 0 for i in input]`), or a state predicate in language (e.g. "Make each element in the input true if it is even") which would then be evaluated by the language model.

Using this experimental paradigm, we can ask many questions pertinent to the question of task decomposition. First, do these subgoals help search at all? It may be the case that, for some domains, language models cannot generate semantically meaningful subgoals without extensive fine-tuning. Second, how much do task descriptions help subgoal generation? While there is evidence that language models can perform program induction [24, 25], providing natural language descriptions of the task-to-solve could allow the language model to generate subgoals even if it could not induce the solution.

Additionally, we can also collect human data to generate a baseline from which to compare these language model-generated subgoals. Like the language model subgoals, these subgoals can take many forms: such as language descriptions of intermediate states or exact intermediate states. Specifically, we can compare the performance of the program synthesis engine on subgoals generated by humans with subgoals generated by language models. For these comparisons, we could run on many domains, such as the Language-annotated Abstraction and Reasoning Corpus (abstract reasoning) [26, 27], CLEVR (visual reasoning) [28], INT (automated theorem proving) [29], or Sokoban (puzzle game) [30].

A stretch goal for this thesis work is to use the findings from Sections 3.1 and 3.2 to motivate a new prompting or decoding technique which improves the performance of language models in generating subgoals. This would be a unique contribution of a new language model prompting or decoding method which was motivated by work in cognitive science; previous prompting techniques are largely motivated by other intuitions and trial-and-error. Even without this final contribution, this experiment would a). provide valuable insight into how we can use language models to improve program synthesis performance and b). provide analyses comparing the capabilities of humans and language models on subgoal generation for complex reasoning tasks.

3.4 Timeline

Sept - Oct. 2023
Development of interface for human studies (decomposition with Theory of Mind and stochastic planning) and preliminary experiments with language model subgoals for program synthesis.

Nov - Dec. 2023
Preliminary experiments with human subjects and development of models of human behavior. Collect human subgoals on complex reasoning domains.

Jan - Feb. 2024
Run experiments to collect human data on both decomposition studies. Compare program synthesis performance between language model-generated subgoals and human-generated subgoals.

March - April. 2024
Compare human performance to model predictions and create a non-linear task decomposition process with language models that is informed by human studies.

May 2024
Write thesis.

References

[1] S. Carey, *The Origin of Concepts*, ser. Oxford series in cognitive development. Oxford University Press, 2011. [Online]. Available: https://books.google.com/books?id=h9J2DwAAQBAJ

[2] J. Fodor and E. Lepore, "The red herring and the pet fish: why concepts still can't be prototypes," *Cognition*, vol. 58, no. 2, pp. 253–270, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/001002779500694X

[3] H. Kamp and B. Partee, "Prototype theory and compositionality," *Cognition*, vol. 57, no. 2, pp. 129–191, 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0010027794006599

[4] E. M. Markman, *Categorization and Naming in Children: Problems of Induction*. The MIT Press, 02 1991. [Online]. Available: https://doi.org/10.7551/mitpress/1750.001.0001

[5] D. N. Osherson and E. E. Smith, "On the adequacy of prototype theory as a theory of concepts," *Cognition*, vol. 9, no. 1, pp. 35–58, 1981. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0010027781900135

[6] E. E. Smith and D. N. Osherson, "Conceptual combination with prototype concepts," *Cognitive Science*, vol. 8, no. 4, pp. 337–361, 1984. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1207/s15516709cog0804_2

[7] S. Piantadosi and R. Aslin, "Compositional reasoning in early childhood," *PLOS ONE*, vol. 11, no. 9, pp. 1–12, 09 2016. [Online]. Available: https://doi.org/10.1371/journal.pone.0147734

[8] R. Anil, A. M. Dai, O. Firat, M. Johnson, D. Lepikhin, A. Passos, S. Shakeri, E. Taropa, P. Bailey, Z. Chen, E. Chu, J. H. Clark, L. E. Shafey, Y. Huang, K. Meier-Hellstern, G. Mishra, E. Moreira, M. Omernick, K. Robinson, S. Ruder, Y. Tay, K. Xiao, Y. Xu, Y. Zhang, G. H. Abrego, J. Ahn, J. Austin, P. Barham, J. Botha, J. Bradbury, S. Brahma, K. Brooks, M. Catasta, Y. Cheng, C. Cherry, C. A. Choquette-Choo, A. Chowdhery, C. Crepy, S. Dave, M. Dehghani,

S. Dev, J. Devlin, M. Díaz, N. Du, E. Dyer, V. Feinberg, F. Feng, V. Fienber, M. Freitag, X. Garcia, S. Gehrmann, L. Gonzalez, G. Gur-Ari, S. Hand, H. Hashemi, L. Hou, J. Howland, A. Hu, J. Hui, J. Hurwitz, M. Isard, A. Ittycheriah, M. Jagielski, W. Jia, K. Kenealy, M. Krikun, S. Kudugunta, C. Lan, K. Lee, B. Lee, E. Li, M. Li, W. Li, Y. Li, J. Li, H. Lim, H. Lin, Z. Liu, F. Liu, M. Maggioni, A. Mahendru, J. Maynez, V. Misra, M. Moussalem, Z. Nado, J. Nham, E. Ni, A. Nystrom, A. Parrish, M. Pellat, M. Polacek, A. Polozov, R. Pope, S. Qiao, E. Reif, B. Richter, P. Riley, A. C. Ros, A. Roy, B. Saeta, R. Samuel, R. Shelby, A. Slone, D. Smilkov, D. R. So, D. Sohn, S. Tokumine, D. Valter, V. Vasudevan, K. Vodrahalli, X. Wang, P. Wang, Z. Wang, T. Wang, J. Wieting, Y. Wu, K. Xu, Y. Xu, L. Xue, P. Yin, J. Yu, Q. Zhang, S. Zheng, C. Zheng, W. Zhou, D. Zhou, S. Petrov, and Y. Wu, "Palm 2 technical report," 2023.

[9] K. Singhal, T. Tu, J. Gottweis, R. Sayres, E. Wulczyn, L. Hou, K. Clark, S. Pfohl, H. Cole-Lewis, D. Neal, M. Schaekermann, A. Wang, M. Amin, S. Lachgar, P. Mansfield, S. Prakash, B. Green, E. Dominowska, B. A. y Arcas, N. Tomasev, Y. Liu, R. Wong, C. Semturs, S. S. Mahdavi, J. Barral, D. Webster, G. S. Corrado, Y. Matias, S. Azizi, A. Karthikesalingam, and V. Natarajan, "Towards expert-level medical question answering with large language models," 2023.

[10] R. Luo, L. Sun, Y. Xia, T. Qin, S. Zhang, H. Poon, and T.-Y. Liu, "BioGPT: generative pre-trained transformer for biomedical text generation and mining," *Briefings in Bioinformatics*, vol. 23, no. 6, sep 2022. [Online]. Available: https://doi.org/10.1093%2Fbib%2Fbbac409

[11] N. Wies, Y. Levine, and A. Shashua, "Sub-task decomposition enables learning in sequence to sequence tasks," 2023.

[12] L. Bertolini, J. Weeds, and D. Weir, "Testing large language models on compositionality and inference with phrase-level adjective-noun entailment," in *Proceedings of the 29th International Conference on Computational Linguistics*. Gyeongju, Republic of Korea: International Committee on Computational Linguistics, Oct. 2022, pp. 4084–4100. [Online]. Available: https://aclanthology.org/2022.coling-1.359

[13] J. Reppert, B. Rachbach, C. George, L. Stebbing, J. Byun, M. Appleton, and A. Stuhlmüller, "Iterated decomposition: Improving science qa by supervising reasoning processes," 2023.

[14] E. Zelikman, Q. Huang, G. Poesia, N. D. Goodman, and N. Haber, "Parsel: A (de-)compositional framework for algorithmic reasoning with language models," 2023.

[15] O. Press, M. Zhang, S. Min, L. Schmidt, N. A. Smith, and M. Lewis, "Measuring and narrowing the compositionality gap in language models," 2022.

[16] H. Hu, H. Lu, H. Zhang, W. Lam, and Y. Zhang, "Chain-of-symbol prompting elicits planning in large langauge models," 2023.

[17] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. Le, and D. Zhou, "Chain-of-thought prompting elicits reasoning in large language models," 2023.

[18] M. Geva, D. Khashabi, E. Segal, T. Khot, D. Roth, and J. Berant, "Did aristotle use a laptop? a question answering benchmark with implicit reasoning strategies," *Transactions of the Association for Computational Linguistics*, vol. 9, pp. 346–361, 2021.

[19] D. Premack and G. Woodruff, "Does the chimpanzee have a theory of mind?" *Behavioral and Brain Sciences*, vol. 1, no. 4, p. 515–526, 1978.

[20] M. Ho, R. Saxe, and F. Cushman, "Planning with theory of mind: Trends in cognitive sciences - cell press," Nov 2022. [Online]. Available: https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(22)00185-1

[21] C. G. Correa, M. K. Ho, F. Callaway, N. D. Daw, and T. L. Griffiths, "Humans decompose tasks by trading off utility and computational cost," 2022.

[22] F. Callaway, B. van Opheusden, S. Gul, P. Das, P. M. Krueger, T. L. Griffiths, and F. Lieder, "Rational use of cognitive resources in human planning," *Nature Human Behaviour*, vol. 6, no. 8, pp. 1112–1125, Aug 2022. [Online]. Available: https://doi.org/10.1038/s41562-022-01332-8

[23] M. Zawalski, M. Tyrolski, K. Czechowski, T. Odrzygóźdź, D. Stachura, P. Piekos, Y. Wu, Łukasz Kuciński, and P. Miłoś, "Fast and precise: Adjusting planning horizon with adaptive subgoal search," 2023.

[24] E. Nijkamp, B. Pang, H. Hayashi, L. Tu, H. Wang, Y. Zhou, S. Savarese, and C. Xiong, "Codegen: An open large language model for code with multi-turn program synthesis," in *The Eleventh International Conference on Learning Representations*, 2023. [Online]. Available: https://openreview.net/forum?id=iaYcJKpY2B_

[25] Z. Yang, L. Dong, X. Du, H. Cheng, E. Cambria, X. Liu, J. Gao, and F. Wei, "Language models as inductive reasoners," 2022.

[26] F. Chollet, "On the measure of intelligence," 2019.

[27] S. Acquaviva, Y. Pu, M. Kryven, T. Sechopoulos, C. Wong, G. E. Ecanow, M. Nye, M. H. Tessler, and J. B. Tenenbaum, "Communicating natural programs to humans and machines," 2023.

[28] J. Johnson, B. Hariharan, L. van der Maaten, L. Fei-Fei, C. L. Zitnick, and R. Girshick, "Clevr: A diagnostic dataset for compositional language and elementary visual reasoning," 2016.

[29] Y. Wu, A. Q. Jiang, J. Ba, and R. Grosse, "Int: An inequality benchmark for evaluating generalization in theorem proving," 2021.

[30] N. Lipovetzky and H. Geffner, "Width and serialization of classical planning problems." vol. 242, 01 2012, pp. 540–545.